

WHAT IS CLAIMED IS:

1. An acoustic model creating method of creating an HMM (Hidden Markov Model) by optimizing, for each state, Gaussian distribution numbers of the respective states constituting the HMM and retraining the optimized HMM using training speech data, the method comprising:

setting plural types of the Gaussian distribution numbers from a predetermined value to a maximum distribution number for each of the plurality of states constituting the HMM;

computing a description length for each of the plurality of states having the plural types of Gaussian distribution numbers using a Minimum Description Length criterion;

selecting a state having the Gaussian distribution number whose description length is minimum, for every state; and

constructing the HMM in accordance with the state having the Gaussian distribution number whose description length is minimum, selected for every state, and retraining the constructed HMM using the training speech data.

2. An acoustic model creating method according to Claim 1, wherein, for the Minimum Description Length criterion, a description length  $li(\chi^N)$  using a model  $i$  when a model set  $\{1, \dots, i, \dots, I\}$  and data  $\chi^N = \{\chi_1, \dots, \chi_N\}$  ( $N$  being a data length) are given is expressed as the following general equation,

$$l_i(x^N) = -\log P_{\theta(i)}(x^N) + \frac{\beta_i}{2} \log N + \log I$$

$\theta(i)$ : parameter of model  $i$

$\theta^{(i)} = \text{maximum likelihood estimate of } \theta_1^{(i)}, \dots, \theta_{\beta_i}^{(i)}$

$\beta_i$ : dimension (degree of freedom) of model  $i$

and in the general equation that computes the description length, the model set  $\{1, \dots, i, \dots, I\}$  is considered as a set of states in which plural types of the Gaussian distribution numbers from a predetermined value to the maximum distribution number are set for a predetermined state in a predetermined HMM, where, when the number of types of the Gaussian distribution numbers is  $I$  ( $I$  is an integer satisfying  $I \geq 2$ ), then  $1, \dots, i, \dots, I$  are symbols that specify the respective distribution number types from a first type to an  $I$ -th type, and the general equation

is used as an equation for computing the description length of the state having an  $i$ -th type of distribution number out of  $1, \dots, i, \dots, I$ .

3. An acoustic model creating method according to Claim 2, in the general equation that computes the description length, the second term on the right side of the equation being multiplied by a weighting coefficient  $\alpha$ .

4. An acoustic model creating method according to Claim 2, in the general equation that computes the description length, the second term on the right side of the equation being multiplied by the weighting coefficient  $\alpha$ , and the third term on the right side being omitted.

5. An acoustic model creating method according to Claim 2, the data  $\chi^N$  being a set of the respective training speech data obtained by matching in time series a plurality of the training speech data with the respective states of the HMMs for every state, using the HMMs in which the respective states have any one of the Gaussian distribution numbers from the predetermined value to the maximum distribution number.

6. An acoustic model creating method according to Claim 5, the any one of the Gaussian distribution numbers being the maximum distribution number.

7. An acoustic model creating method according to Claim 1, the HMMs being syllable HMMs.

8. An acoustic model creating method according to Claim 7, wherein, for a plurality of syllable HMMs having a same consonant or a same vowel in the syllable HMMs, the syllable HMMs having the same consonant out of the states constituting the syllable HMMs tie an initial state or at least two states including an initial state in the syllable HMMs, and the syllable HMMs having the same vowel tie a final state of the states having self loops or at least two states including the final state in the syllable HMMs.

9. A speech recognition device that recognizes input speech using HMMs (Hidden Markov Models) as acoustic models for feature data obtained by feature analysis of the input speech,

the HMMs created by the acoustic model creating method according to Claim 1 being used as the HMMs which are the acoustic models.